Module de Pobabilités 2
Chapitre 3 : Processus de Poisson
Séance 8

Responsable du cours: Dr. **Metiri Farouk,**
Université de Badji Mokhtar -Annaba-

Mail address: fmetiri@yahoo.fr

Chapitre 3: Le Processus de Poisson

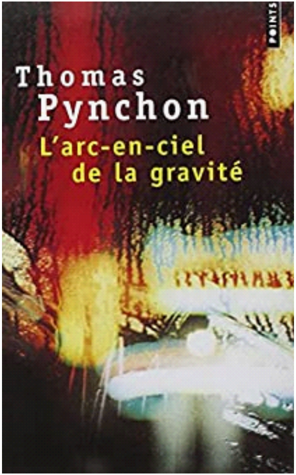
Section 3.1: Rappels et Notions de Base Sur La Loi de Poisson:

Une variable aléatoire discrète X suit une loi de Poisson de paramètre $\lambda \in \Re^{+}$, ce qu'on le note $X \sim P(\lambda)$, si $X(\Omega) = \mathbb{N}$
et si $\forall x \in \mathbb{N}$,

$$P[X = x] = \frac{\lambda^x}{x!} e^{-\lambda}$$

Cette loi est en général utilisée des événements rares comme le nombre d'accidents de voiture, le nombre de mutations génétiques fixées dans l'ADN,. . .En fait, la loi de Poisson a été introduite en 1838 par Siméon-Denis Poisson dans son ouvrage "Recherches sur la probabilité des jugements en matière criminelle et en matiére civile".

Même en littérature, Dans le roman de **Thomas Pynchon**, **"L'Arc-en-ciel de la gravité"**, un des personnages, le statisticien Roger Mexico, utilise la loi de Poisson pour cartographier les zones d'impact des fusées allemandes V_2 sur la ville de Londres durant la Seconde Guerre mondiale.



Les graphiques représentés aux figures 3.1, 3.2 et 3.3 permettent de voir l'impact du paramètre λ sur la loi de la variable.

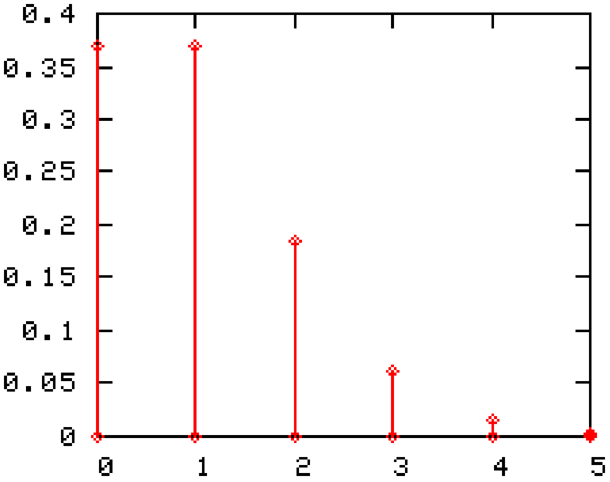


Figure: 3.1: diagramme en bâtons d'une loi de Poisson de paramètre $\lambda = 1 : P(X = k) = \frac{1^k}{k!} e^{-1}, K \in \mathbb{N}$.

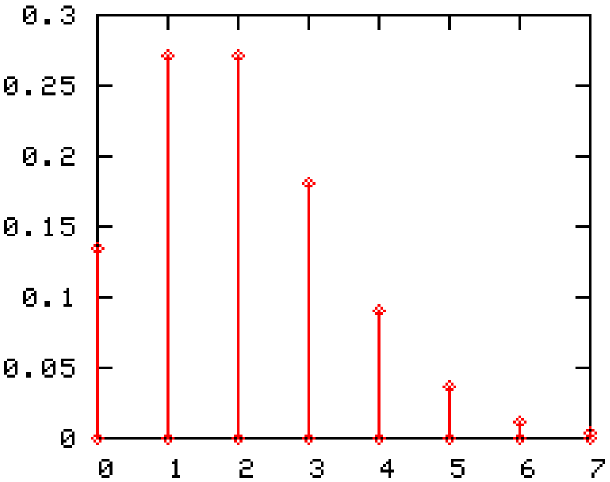


Figure: 3.2: diagramme en bâtons d'une loi de Poisson de paramètre $\lambda = 2 : P(X = k) = \frac{2^k}{k!}e^{-2}$, $K \in \mathbb{N}$.

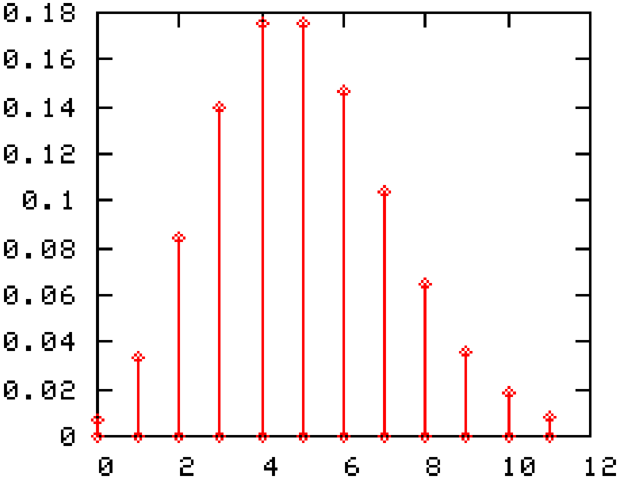


Figure: 3.3: diagramme en bâtons d'une loi de Poisson de paramètre $\lambda = 5 : P(X = k) = \frac{5^k}{k!} e^{-5}$, $K \in \mathbb{N}$.

Proposition 3.1: La fonction génératrice des probabilités d'une variable aléatoire discrète distribuée selon une loi de Poisson de paramètre λ est donnée par:

$$\Phi_X(t) = E\left[e^{tX}\right] = e^{\lambda(e^t - 1)}$$

Preuve:

$$\Phi_X(t) = E\left[e^{tX}\right] = \sum_{k=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^t\right)^k}{k!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

$$\Phi'_X(t) = \lambda e^t e^{\lambda(e^t - 1)}, \text{ donc } E[X] = \Phi'_X(0) = \lambda$$

$$\Phi''_X(t) = (1 + \lambda e^t)\lambda e^t e^{\lambda(e^t - 1)}, \text{donc } E\left[X^2\right] = \Phi''_X(0) = \lambda + \lambda^2$$

D'où

$$V[X] = E\left[X^2\right] - (E[X])^2 = \Phi''_X(0) - (\Phi'_X(0))^2 = \left(\lambda + \lambda^2\right) - (\lambda)^2 = \lambda$$

Proposition 3.2: La somme de n variables indépendantes de loi

Poisson suit encore une loi de Poisson : soient $X_1, ..., X_n$ des variables indépendantes de loi respective $P(\lambda_i)$, alors:

$$\sum_{i=1}^{n} X_i \sim P(\sum_{i=1}^{n} \lambda_i)$$

Preuve: On considère $Z = \sum_{i=1}^{n} X_i$ et On calcule sa fonction génératrice des moments:

$$\Phi_Z(t) = E\left[e^{tZ}\right] = E\left[e^{t\sum_{i=1}^{n} X_i}\right] = E\left[e^{tX_1+tX_2+...+tX_n}\right]$$

En utilisant le fait que la fonction génératrice caractérise la loi et sachant que les X_i sont indépendantes, on obtient:

$$\begin{aligned}\Phi_Z(t) &= E\left[e^{tX_1}\right]E\left[e^{tX_2}\right]...E\left[e^{tX_n}\right] = e^{\lambda_1(e^t-1)}.e^{\lambda_2(e^t-1)}.....e^{\lambda_n(e^t-1)} \\ &= e^{\sum_{i=1}^{n}\lambda_i(e^t-1)} : \text{ fct génératrice d'une v.a} \sim P(\sum_{i=1}^{n}\lambda_i)\end{aligned}$$

ce qui conclut la preuve.

L'espérance conditionnelle:

Pour de nombreux problèmes concrets (prédiction, observation incomplète, etc.) il est important de pouvoir estimer une variable aléatoire sur laquelle on n'a qu'une information partielle. Dès lors, on comprend l'importance de la notion d'espérance conditionnelle.

Considérons deux variables aléatoires X et Y . On appelle espérance conditionnelle de X sachant Y , notée $E[X|Y]$, la variable aléatoire qui prend la valeur $E[X|Y=y]$ en y (C'est la variable aléatoire qui donne la valeur moyenne de X quand on connait Y).

Cette quantité est une fonction de Y donc une variable aléatoire $\sigma(Y) - mesurable$ (et ne dépend surtout pas de X).

Si X et Y sont des variables *indépendantes* alors l'espérance conditionnelle est *constante*:
L'espérance conditionnelle vérifie la relation suivante:

$$E[X] = E[E[X|Y]].$$

Si Y est discrète, l'expression précédente se réécrit:

$$E[X] = \sum_y E[X|Y = y]P(Y = y)$$

et si Y est une variable continue de densité $f_Y(y)$ on a:

$$E[X] = \int_{-\infty}^{+\infty} E[X|Y = y]f_Y(y)dy$$

Somme d'un nombre aléatoire de variables aléatoires:
On pose N le nombre d'accidents de voitures pendant une année. Pour le $i^{ème}$ accident, on note X_i le montant des indemnités que la compagnie d'assurance verse au conducteur. On suppose que les X_i sont indépendants et de même loi et indépendants de N. Le montant total S que la compagnie devra débourser au cours d'une année est:

$$S = \sum_{i=1}^{N} X_i$$

On souhaite calculer sa valeur moyenne. Comme X s'écrit à l'aide de N, il est naturel de conditionner par rapport à cette valeur. On a

$$E[S] = E\left[E[\sum_{i=1}^{N} X_i|N]\right]$$

Par ailleurs, on a

$$\begin{aligned} E[\sum_{i=1}^{N} X_i|N &= n] = \sum_{i=1}^{N} E[X_i] \text{ par indépendance entre les } X_i \text{ et } N \\ &= NE\left[X\right] \text{car les } X_i \text{ sont de même loi.} \end{aligned}$$

Par conséquent $E[S|N] = NE[X]$, on en déduit que

$$E[S] = E\left[E[\sum_{i=1}^{N} X_i|N]\right] = E\left[NE[X]\right] = E\left[N\right]E\left[X\right]$$

On peut généraliser le calcul d'espérance au calcul de l'espérance de n'importe quelle fonction g de X : si
Y est discrète, on a

$$E\left[g(X)\right] = \sum_{y} E[g(X)|Y = y]P\left(Y = y\right)$$

et si Y est une variable continue de densité $f_Y(y)$ on a:
$E\left[g(X)\right] = \int_{-\infty}^{+\infty} E[g(X)|Y = y]f_Y(y)dy$

Ceci est utile pour calculer la variance ou la fonction génératrice d'une variable aléatoire. Notamment pour la variance $Var(X) = E[X^2] - (E[X])^2$ on obtient la formule suivante:

Formule de décomposition de la variance:

$$Var(X) = E[Var(X \mid Y)] + Var[E[X \mid Y]]$$

Preuve: On a

$$E[Var(X \mid Y)] = E\left[E\left[X^2 \mid Y\right] - E[X \mid Y]^2\right]$$

$$= E\left[E\left[X^2 \mid Y\right]\right] - E\left[E[X \mid Y]^2\right] = E\left[X^2\right] - E\left[E[X \mid Y]^2\right]$$

et

$$Var[E[X \mid Y]] = E\left[E[X \mid Y]^2\right] - [E[E[X \mid Y]]]^2 = E\left[E[X \mid Y]^2\right] - E[X]^2$$

En sommant les deux expressions, on obtient le résultat

- On reprend l'exemple précédent et on veut calculer la variance des remboursements au cours d'une année.
On va utiliser la formule de décomposition de la variance:

$$Var(S) = E[Var(S \mid N)] + Var\left[E\left[S \mid N\right]\right]$$

On a $E[S|N] = NE[X]$, par conséquent

$$Var\left[E\left[S \mid N\right]\right] = Var\left[NE[X]\right] = E\left[X\right]^2 Var\left(N\right) \tag{3.1}$$

D'un autre coté, $E[Var(S \mid N)] = ?$
On a

$$Var(S \mid N) = Var(\sum_{i=1}^{N} X_i|N) = \sum_{i=1}^{N} Var(X_i|N)$$

car les X_i sont indépendants.

$$= \sum_{i=1}^{N} Var(X_i) \text{ car les } X_i \text{ et } N \text{ sont indépendants.}$$

Par la formule de décomposition de la variance, on en déduit que

$$Var(S) = E[X]^2 Var(N) + Var(X)E[N]. \tag{3.3}$$

Notamment, si on suppose que le nombre de sinistres suit une loi de Poisson $P(\lambda)$, on a $E[N] = Var(N) = \lambda$
et donc

$$Var(S) = \lambda(E[X]^2 + Var(X)) = \lambda E[X^2].$$

Des exemples précédents on obtient la proposition suivante.

Proposition: Espérance et variance d'une somme aléatoire de variables aléatoires:

On considère une suite $(X_i)i_{\geq 1}$ des variables indépendantes et de même loi que X et N une variable aléatoire. On suppose que N et les X_i sont indépendants, alors:

$$E[\sum_{i=1}^{N} X_i] = E\left[N\right] E[X]$$

$$Var\left[\sum_{i=1}^{N} X_i\right] = Var\left[NE[X]\right] = E\left[X\right]^2 Var\left(N\right) + Var(X)E\left[N\right]$$

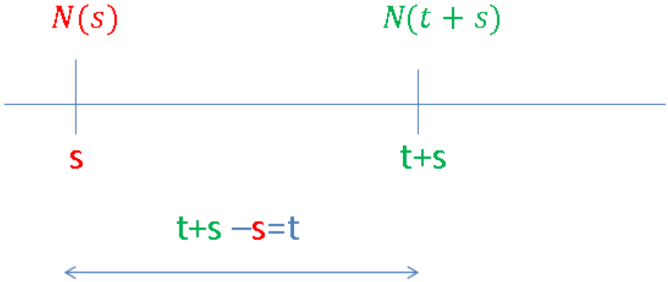
Section3.2: Processus de Poisson homogène

Definition

Un processus de comptage $\{N_t; t \geq 0\}$à valeurs entières est un processus de Poisson de paramètre
$\lambda > 0$ si:
$i)$ (N_t) est un processus de comptage à accroissements indépendants (c'est à dire pour $t > s$, le
nombre de sauts $N(t) - N(s)$intervenus sur $[s, t]$ est indépendant du nombre de sauts $N(s)$ intervenus
avant l'instant s.
$ii)$ (N_t) est un processus de comptage à accroissements stationnaires (ayant tous la même loi)
$iii)$ le nombre d'événements sur un intervalle de longueur t est une variable aléatoire qui suit la loi de
Poisson de paramètre λt :

$$\forall s, t \text{ et } n \geq 0, P\left(N\left(t+s\right) - N(s) = n\right) = \frac{\left(\lambda t\right)^n}{n!} e^{-\lambda t}.$$

le nombre moyen de sauts par intervalle de temps
= Le paramètre de la loi de poisson



- Pour un processus de comptage à accroissements stationnaires (comme le processsu de Poisson), la loi du nombre de sauts dans un intervalle de temps donné est indépendante de tout le passé du processus, jusqu'au premier instant de l'intervalle de temps.
- Comme la variable aléatoire $N(t) \sim P(\lambda t)$, on remarque que **le nombre moyen de sauts par intervalle de temps** est:

$$\frac{N(t)}{t} = \frac{\lambda t}{t} = \lambda$$

Remark

Remarque: *Soit* $(N(t))_{t\geq 0}$ *un processus de Poisson d'intensité* λ *. Le nombre de sauts sur un petit intervalle de temps ne dépasse généralement pas* 1*. Plus précisément, soit* $h > 0$*, pour* h *suffisamment petit on a* $\lambda h < 1$ *et donc:*

$$\begin{aligned} P\left(N(h) = 0\right) &= \frac{\lambda^0}{0!} e^{-\lambda h} = e^{-\lambda h}; P\left(N(h) = 1\right) = \lambda h e^{-\lambda h}; \\ P\left(N(h) \geq 2\right) &\leq \frac{(\lambda h)^2}{2} e^{-\lambda h}. \end{aligned}$$

et quand h *tend vers* 0*, on a*

$$\frac{P\left(N(h) = 1\right)}{h} \longrightarrow \lambda \ et \ \frac{P\left(N(h) \geq 2\right)}{h} \longrightarrow 0$$

Distribution des temps d'attente et des inter-arrivées:
Un processus de Poisson sert à compter des sauts (ou des événements). On peut se demander à quels moments arrivent ces sauts, et quelle distance il y a entre deux sauts. Les sauts arrivent de façon aléatoire, on va donc essayer de trouver la **loi d'attente entre deux sauts**: **la loi des instants inter-arrivées**.

Considérons un processus de Poisson $(N(t))_{t\geq 0}$. Le premier saut arrive à un instant aléatoire T_1, puis il faudra attendre un temps T_2 avant que le second saut survienne, puis un temps T_3 ainsi de suite. . . .
On note T_n le temps écoulé entre le $(n-1)^{ème}$ saut et le $n^{ème}$ saut.
La suite des instants $(T_n)_{n\geq 1}$ est appelée suite des instants inter-arrivées.
Cherchons la loi de T_n.

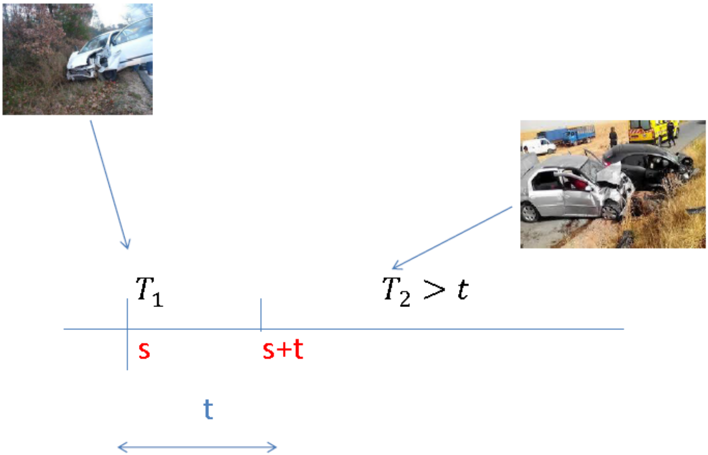
La suite des instants inter-arrivées $(T_n)_{n\geq 1}$ est une suite de variables indépendantes identiquement distribuées de loi exponentielle $exp(\lambda)$. Par conséquent, **la durée moyenne d'un intervalle de temps entre deux sauts est** $\frac{1}{\lambda}$.

Étudions le premier instant de saut T_1. Si à l'instant t on a $N(t) = 0$ ceci signifie que l'on n'a pas encore sauté et donc $T_1 > t$. Par conséquent:

$$P\left(T_1 > t\right) = P\left(N(t) = 0\right) = \frac{\lambda^0}{0!}e^{-\lambda t} = e^{-\lambda t}$$

Ceci implique que T_1 suit la loi exponentielle de paramètre λ.

Maintenant étudions la loi de T_2 en conditionnant par rapport à T_1 : soit $t, s > 0$



$$P\left(T_2 > t \mid T_1 = s\right) = P\left(\text{pas de sauts sur l'intervalle}]s, s+t] \mid T_1 = s\right)$$

Le nombre de sauts sur l'intervalle $]s, s+t]$ suit une loi de Poisson $P(\lambda t)$ et il est indépendant de ce qui s'est passé avant, par conséquent

$$\begin{aligned} P\left(T_2 > t \mid T_1 = s\right) &= P\left(\text{pas de saut sur l'intervalle}]s, s+t]\right) \\ &= \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t} \\ &= 1 - F_{T_2}(t) \end{aligned}$$

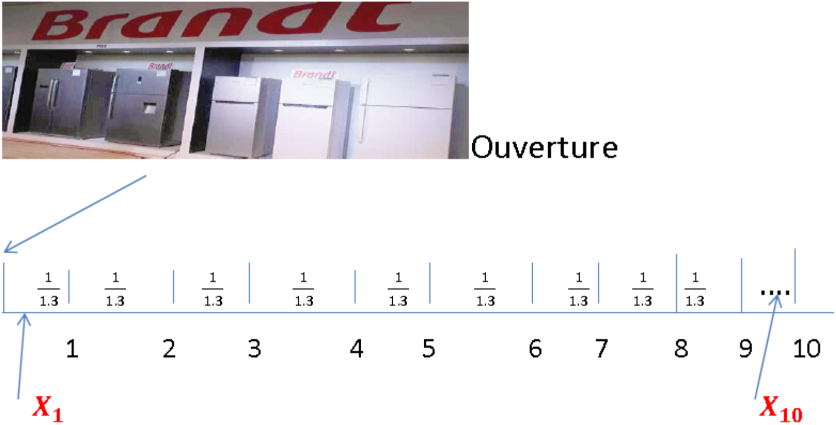
Par conséquent $T_2 \sim exp(\lambda)$ et les variables T_2 et T_1 sont indépendantes. En **itérant** ce raisonnement, on prouve que le **loi d'attente entre deux sauts (la loi des instants inter-arrivées)** est la *loi exponentielle*.

Exemple 3.1: On suppose que les clients arrivent au magasin d'appareils électroménagers selon un processus de Poisson d'intensité $\lambda = 1,3$ par heure.

- Le temps moyen avant de voir arriver le $9^{ème}$ client est égal à $E[X_9] = 9.E[X] = 9.\frac{1}{\lambda} = \frac{9}{1,3} = 6,92$ heures.
- Si maintenant on regarde la probabilité que le temps d'attente entre le $9^{ème}$ client et le $10^{ème}$

client soit supérieur à 90 minutes :

$$
\begin{aligned}
P(X_{10} > 90 minutes) &= P(X_{10} > 1,5 \text{ heures}) \\
&= 1 - P(X_{10} \leq 1,5 \text{ heures}) \\
&= 1 - F_X(1,5) = e^{-\lambda(1,5)} = e^{-1,3(1,5)} = 0,1423
\end{aligned}
$$


Brandt
Ouverture
1/1.3
1/1.3
1/1.3
1/1.3
1/1.3
1/1.3
1/1.3
1/1.3
1/1.3
....
1 2 3 4 5 6 7 8 9 10
X_1
X_{10}

Exercice 3.1: Les questions sont indépendantes:

(1) Un actuaire estime que le nombre N de réclamations suit une loi de Poisson.

- Trouver $Var\,[N]$ sachant que $\frac{P(N=2)}{P(N=4)} = 3.$

(2) Soit X_1, X_2, X_3, X_4, X_5 des v.a indépendantes toutes de loi de poisson de paramètres respectifs $3, 1, 2, 1$ et 4.

- Trouver la probabilité que $X_1 + X_2 + X_3 + X_4 + X_5$ soit au plus 2.

(3) Un actuaire constate que la probabilité qu'un assuré n'ait aucun accident est dix fois plus grande que celle d'en avoir au moins un durant l'année. En supposant que le nombre d'accidents de l'assuré suit une loi de Poisson:

- Trouver la probabilité que l'assuré air exactement deux accidents durant l'année.

(4) Soit X le nombre hebdomadaire d'accidents dans une petite ville. On suppose que X suit une loi de Poisson de moyenne $3.$

- Trouver la probabilité qu'il y ait un seul accident durant les deux prochaines semaines.

Exercice 3.2:

Dans un centre d'appel téléphonique, les appels arrivent selon un processus de Poisson de taux 10 appels par heure.

1. Si une téléphoniste fait une pause de 10 heures a $10h30$, combien d'appels va-t-elle rater en moyenne durant sa pause?

2. Quelle est la probabilité qu'elle a raté au plus 2 appels?

3.Sachant que 4 appels arrivent entre 10 heures et 11 heures, quelle est la probabilité qu'elle n'a raté aucun appel durant sa pause? Qu'elle n'a raté qu'un appel?

4. Sachant qu'il y aura 2 appels entre $10h30$ et 11 heures, quelle est la probabilité qu'ils arrivent tous entre $10h30$ et $10h45$?.

Solution 3.1:

(1) On sait que $Var[N]$ =le paramètre de loi de poisson $=\lambda$.

$$
\begin{aligned}
Var[N] &= \lambda = \frac{\frac{\lambda^2}{2!}e^{-\lambda}}{\frac{\lambda^4}{4!}e^{-\lambda}} = 3 \Longrightarrow \lambda^2 = \frac{12}{3} = 4 \\
&\Longrightarrow \lambda = \left\{ \begin{array}{c} -2 \text{ impossible} \\ 2 \text{ acceptée} \end{array} \right. = Var[N]
\end{aligned}
$$

(2) Selon la propriétés de la loi de Poisson, la v.a $X = X_1 + X_2 + X_3 + X_4 + X_5$ suit la loi de Poisson d'intensité $3 + 1 + 2 + 1 + 4 = 11$. $X \sim P(11)$.

La probabilité que X soit au plus 2 est:

$$
\begin{aligned}
P(X \leq 2) &= P(X=0) + P(X=1) + P(X=2) \\
&= \frac{11^0}{0!}e^{-11} + \frac{11^1}{1!}e^{-11} + \frac{11^2}{2!}e^{-11} = \left(1 + 11 + \frac{121}{2}\right)e^{-11} \\
&= \frac{145}{2}e^{-11}
\end{aligned}
$$

(3) Soit $N \sim P(\lambda)$.On a:

$$P(N=0) = 10P(N \geq 1) = 10(1 - P(N < 1))$$

$$\Longrightarrow \quad \frac{\lambda^0}{0!}e^{-\lambda} = 10(1 - P(N=0)) \Longrightarrow e^{-\lambda} = 10\left(1 - e^{-\lambda}\right)$$

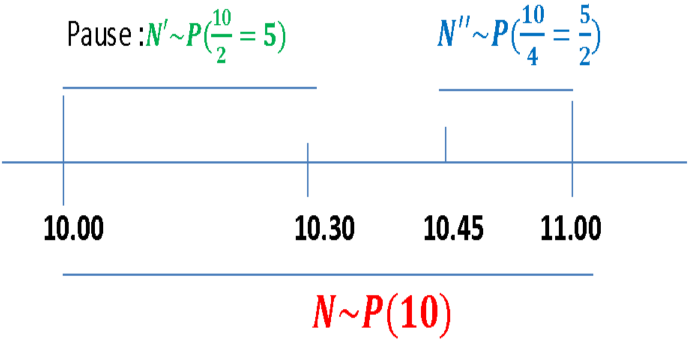
$$\lambda \quad = \quad -\ln\left(\frac{10}{11}\right) = 0.0953$$

Donc,

$$P(N=2) = \frac{0.0953^2}{2!}e^{-0.0953} = 0.0041$$

(4) On a $X \sim P(3)$.
durant un semaine, il y a $E[X] = \lambda = 3$ accidents en moyenne, ca signifie que durant deux semaines il y a en moyenne $E[X] = \lambda' = \lambda t = 3.2 = 6$ accidents.
La probabilité qu'il y ait un seul accident durant deux semaines est $P(X=1) = \frac{6^1}{1!}e^{-6} = 0.0149$

Solution 3.2:
On considère que le nombre d'appels téléphoniques est une v.a $N \sim P(10)$.



(1) En moyenne, il y a $\lambda = E[N] = 10$ appels chaque heure. Donc, dans une demi heure, il y a en moyenne 5 appels ratés car le nombre d'appels téléphoniques ratés est une v.a $N'_{10-10.30} \sim P(5)$.
La téléphoniste va rater en moyenne: $5appels = E\left[N^1_{10-10.30}\right]$.
(2) La probabilité qu'elle a raté au plus 2 appels est:

$$
\begin{aligned}
& P\left(N'_{10-10.30} \leq 2\right) \\
= \; & P\left(N'_{10-10.30} = 0\right) + P\left(N'_{10-10.30} = 1\right) + P\left(N'_{10-10.30} = 2\right) \\
= \; & \frac{5^0}{0!}e^{-5} + \frac{5^1}{1!}e^{-5} + \frac{5^2}{2!}e^{-5} = \frac{37}{2}e^{-5}
\end{aligned}
$$

$$
\begin{aligned}
P\left(N'_{10-10.30}=0 \mid N=4\right) &= \frac{P\left(\left(N'_{10-10.30}=0\right)\ \text{et}\ \left(N'_{10.30-11}=4\right)\right)}{P(N=4)} \\
&= \frac{P\left(N'_{10-10.30}=0\right).P\left(N'_{10.30-11}=4\right)}{\frac{10^4}{4!}e^{-10}} \\
&= \frac{\frac{5^0}{0!}e^{-5}.\frac{5^4}{4!}e^{-5}}{\frac{10^4}{4!}e^{-10}} = \left(\frac{1}{2}\right)^4 = \frac{1}{16}
\end{aligned}
$$

$$
\begin{aligned}
P\left(N'_{10-10.30}=1 \mid N=4\right) &= \frac{P\left(\left(N'_{10-10.30}=1\right)\ \text{et}\ \left(N''_{10.30-11}=3\right)\right)}{P(N=4)} \\
&= \frac{P\left(N'_{10-10.30}=1\right).P\left(N''_{10.30-11}=3\right)}{\frac{10^4}{4!}e^{-10}} \\
&= \frac{\frac{5^1}{1!}e^{-5}.\frac{5^3}{3!}e^{-5}}{\frac{10^4}{4!}e^{-10}} = \frac{1}{4}
\end{aligned}
$$

(4)

$$
\begin{aligned}
& P\left(N^{''}_{10.30-10.45}=2 \mid N^{'}_{10.30-11}=2\right) \\
= \quad & \frac{P\left(\left(N^{''}_{10.30-10.45}=2\right) \text{ et } \left(N''_{10.45-11}=0\right)\right)}{P(N_{10.30-11}=2)} \\
= \quad & \frac{P\left(N^{''}_{10.30-10.45}=2\right).P\left(N''_{10.45-11}=0\right)}{\frac{5^2}{2!}e^{-2}} \\
= \quad & \frac{\frac{\left(\frac{5}{2}\right)^2}{2!}e^{-\frac{5}{2}}.\frac{\left(\frac{5}{2}\right)^0}{0!}e^{-\frac{5}{2}}}{\frac{5^2}{2!}e^{-5}} = \frac{1}{4}.
\end{aligned}
$$

Bon courage